

**OceanSITES:
Global Data Assembly Center**

Bill Burnett

NOAA National Data Buoy Center, Stennis Space Center MS

1. INTRODUCTION

OceanSITES is the international project working towards the coordination and implementation of a global system of sustained multi-disciplinary timeseries observatories. Timeseries fill a unique gap in the sampling provided by other elements of the global ocean observing system, enabling co-located observations of many variables and processes in strategic or representative locations over long periods of time, with high temporal resolution, from (and including) the ocean surface to the seafloor.

The scientific applications of such data are to monitor, detect, understand, and predict changes and related processes in the physical climate state of the ocean, the carbon cycle, and the ecosystem. Operational applications include detection of events, initialization and validation of assimilation products, delivery of constraints or reference data for forecasts (especially biogeochemical and ecosystem relevant ones). In addition there are a variety of technical applications, such as calibration and validation of data and products from other observing system elements.

OceanSITES, through its international steering team, has developed a rationale for timeseries observations and for needing a coordinated global network, and has defined a pilot project consistent with the needs and expectations of the sponsoring bodies GOOS, CLIVAR, and POGO. A major requirement for sites in the project is an open data policy. A global timeseries data management system is under construction via a subgroup of the OceanSITES steering team, including a data format coherent with other past and present efforts.

The in situ, time series-based OceanSITES program represents the logical next step in completing the Global Ocean Observing System. As such, the program now is an official component of the global system organized under JCOMM, and is also one of its action groups under DBCP. Much of the technology is available and many elements are in place already. The main challenge is coordination and assuring sustainability of the system, via common advocacy, recruiting a user base, and sharing the operation among communities and countries.

1.1. National Data Buoy Center (NDBC) GDAC Status

Starting in 2000, NDBC began obtaining and distributing observations from “partners.” These partners are designated as U.S. Integrated Ocean Observing System (IOOS) data providers. NDBC receives these marine meteorological, oceanographic (physical) and water quality observations in real-time, quality controls the observations and distributes the data via the Global Telecommunications System (GTS)/web services. NDBC also serves as the Data Assembly Center (DAC) for the Tropical Atmosphere Ocean (TAO) Pacific array and the tsunameter array which covers the Pacific, Atlantic and Gulf of Mexico. NDBC also quality controls and

maintains data from 60 oil and gas platforms located in the Gulf of Mexico. Thus, NDBC is well suited to serve as a Global Data Assembly Center (GDAC) for OceanSITES, as well as an OceanSITES DAC.

NDBC supports these ~700 platforms by collecting, quality controlling and disseminating the observations in real-time and in delayed mode. Using the NDBC Observing System Monitoring Center (OSMC), OPeNDAP servers and ftp site – NDBC will act as a DAC for physical observations (marine weather and oceanographic – and possibly for biogeochemical variables) for a number of PIs in the United States. NDBC will also serve as the second OceanSITES GDAC and synchronize their OceanSITES files with Coriolis. NDBC proposed a form, similar to the form used to maintain the metadata from the 60 oil and gas platforms, to help maintain all the OceanSITES platforms.

1.2. Tasks proposed

This document describes three tasks in support of OceanSITES, enabling rapid and efficient progress on the data management, quality control, and data dissemination, as well as initiating work towards the generation of products and indicators resulting from this global timeseries network.

2. PROGRESS

The following milestones track the progress made on OceanSITES during fiscal year 2008.

Establishment of an OceanSITES GDAC at NDBC

Overview

The international CLIVAR/GOOS/JCOMM sponsors a coordination project called OceanSITES, a global network of ocean timeseries (or reference) sites located around the world's oceans. The NOAA/NDBC Integrated Ocean Observing System Data Assembly Center (IOOS DAC) and OceanSITES have agreed to make the NDBC IOOS DAC a Global DAC (GDAC), providing a shared and more secure capability together with the Ifremer/Coriolis GDAC in France. These GDACs will provide quality assurance/quality control, provide virtual access to the data, maintain a global timeseries dataset and synchronize catalogues on a periodic basis. The definition and structure of the data system is shown below.

Today NDBC is meeting the U.S. NWS's need for data in the marine environment by assembling and quality checking data from other U.S. government agencies, academia and industry as part of the U.S. Integrated Ocean Observing System (IOOS), a component of the Global Ocean Observing System (GOOS) and the Global Earth Observation System of Systems (GEOSS). The NDBC IOOS DAC is already providing real-time quality control of observations from Woods Hole OceanSites platforms, http://www.ndbc.noaa.gov/station_page.php?station=32st0, as well as operating the Tropical Ocean Atmosphere (TAO) Data Assembly Center. The NDBC IOOS DAC website provides a one-stop shop for real-time and delayed mode, quality controlled

observations for 160 NDBC platforms, 190 NOS water level stations, 220 partner platforms, 55 TAO buoys, 25 NERRS water quality stations and 39 DART tsunami stations.

3. 2008 NDBC ACCOMPLISHMENTS

- **Coordination:** NDBC coordinated with Coriolis personnel to define equipment, hardware, software and telecommunication requirements for the secondary GDAC.
 - Attended the OceanSITES Data Management Team Meeting and the Science Team Meeting in Vienna, Austria - April, 2008.
 - Briefed NDBC's Data Assembly Center and contribution to the Integrated Ocean Observing System (IOOS).
 - Proposed alternative method towards disseminating OceanSITES observations (via OPeNDAP vice FTP).
 - Proposed new metadata sheet (derived from Gulf of Mexico Oil and Gas platform metadata sheet) to be used as the metadata sheet that validates all OceanSITES platforms.
 - Collaborated with Coriolis and other Data Management Members on the format for data distribution.
 - Updated the OceanSITES Users Manual.
 - Coordinated teleconference calls with Coriolis to discuss OceanSITES data management team structure and dissemination of OceanSITES data.
 - Numerous e-mails with Coriolis and Data Management Team personnel on format and dissemination of OceanSITES data.
 - Obtained primary server for the OceanSITES data files.
 - Defined communication requirements to share OceanSITES data with other GDAC on a daily basis.
 - Proposed and agreed to method for sharing files with Coriolis.
 - Worked with Coriolis and SIO to update the DAC format checker.
 - Wrote and edited the 2008 Data Management Team Report
- **Quality Control:** NDBC IOOS DAC continued to integrate additional OceanSITES into their IOOS infrastructure. Observations approved by the Principle Investigator (PI) were released through to the GTS and made available via the OceanSITES ftp server and OPeNDAP server.
 - Developed a local OceanSITES "server" according to specifications approved by the data management group
 - Serves as the DAC for the four Woods Hole Stations
 - Serves as the DAC for the four TAO flux sites
 - Serves as the DAC for the two Scripps sites – and possibly more if wave observations are included.
- **Observing System Monitoring Center (OSMC):** NDBC coordinated efforts with the OSMC project to ensure that the OSMC provides OceanSITES observations and leverages the OPeNDAP/DODS services developed through the OSMC framework.

NDBC will additionally provide all OceanSITES observations through their OPeNDAP server in the approved OceanSITES netCDF formats.

- Continued to update the OSMC Oracle database and NetCDF files with data from the following sources:
 - GODAE data files (Met and Profiles)
 - GTS feed from NWSTG (over 4,800 unique GTS headers)
 - JCOMOPS WMO Platform Cross-Reference (meta-data)
 - NDBC Platform Data
- Worked with Coriolis and SIO to identify enhancements which includes the following:
 - Support for OceanSITES platform operational status information
 - Multiple URLs related to a platform
 - Storage of GTS header/type information
 - Storage of Quality Assurance data
 - Sensor/Instrument information
 - Add Program information, so that platform types are not used incorrectly
 - Address duplicate resolution via an updated view
- **Standardization:** FTP servers at NDBC and Coriolis are online and synchronize available data daily. This data is available from both NDBC and Coriolis. Both GDAC's will develop similar portals to implement viewing services for the global timeseries datasets.
 - Implemented new schema to conduct daily checks with Coriolis GDAC server to update new GDAC observations and format information.
 - Began replicating OceanSITES platforms with Coriolis (GDAC to GDAC) in July 2008 – via <ftp://data.ndbc.noaa.gov/data/oceansites> and <ftp://ftp.ifremer.fr/ifremer/oceansites>
 - Finalized the OceanSITES Global Data Assembly Center form for collecting and maintaining information for OceanSITES data access and data distribution in October 2008.
 - Working with primary OceanSITES PIs to begin updating the metadata forms and will post information on OceanSITES website.